

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/681,878B
Source: 1Fw/b
Date Processed by STIC: 7/5/06

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 07/05/2006

PATENT APPLICATION: US/10/681,878B

TIME: 14:22:51

Input Set : A:\44463336.APP

Output Set: N:\CRF4\07052006\J681878B.raw

3 <110> APPLICANT: CHIANG, VINCENT L.
 4 CARRAWAY, DANIEL T.
 5 SMELTZER, RICHARD H.
 7 <120> TITLE OF INVENTION: PRODUCTION OF SYRINGYL LIGNIN IN GYMNOSPERMS
 9 <130> FILE REFERENCE: 044463-0336
 11 <140> CURRENT APPLICATION NUMBER: 10/681,878B
 12 <141> CURRENT FILING DATE: 2003-10-09
 14 <150> PRIOR APPLICATION NUMBER: 09/796,256
 15 <151> PRIOR FILING DATE: 2001-02-28
 17 <150> PRIOR APPLICATION NUMBER: 08/991,677
 18 <151> PRIOR FILING DATE: 1997-12-16
 20 <150> PRIOR APPLICATION NUMBER: 60/033,381
 21 <151> PRIOR FILING DATE: 1996-12-16
 23 <160> NUMBER OF SEQ ID NOS: 24
 25 <170> SOFTWARE: PatentIn Ver. 3.3
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 1708
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Liquidambar styraciflua
 32 <220> FEATURE:
 33 <221> NAME/KEY: CDS
 34 <222> LOCATION: (48)..(1571)
 36 <400> SEQUENCE: 1

P.6

37	cggcacgagg aaaccctaaa actcacctct cttacccttt ctcttca atg gct ttc	56
38		Met Ala Phe
39		1
41	ctt cta ata ccc atc tca ata atc ttc atc gtc tta gct tac cag ctc	104
42	Leu Leu Ile Pro Ile Ser Ile Ile Phe Ile Val Leu Ala Tyr Gln Leu	
43	5 10 15	
45	tat caa cgg ctc aga ttt aag ctc cca ccc ggc cca cgt cca tgg ccg	152
46	Tyr Gln Arg Leu Arg Phe Lys Leu Pro Pro Gly Pro Arg Pro Trp Pro	
47	20 25 30 35	
49	atc gtc gga aac ctt tac gac ata aaa ccg gtg agg ttc cgg tgt ttc	200
50	Ile Val Gly Asn Leu Tyr Asp Ile Lys Pro Val Arg Phe Arg Cys Phe	
51	40 45 50	
53	gcc gag tgg tca caa gcg tac ggt ccg atc ata tcg gtg tgg ttc ggt	248
54	Ala Glu Trp Ser Gln Ala Tyr Gly Pro Ile Ile Ser Val Trp Phe Gly	
55	55 60 65	
57	tca acg ttg aat gtg atc gta tcg aat tcg gaa ttg gct aag gaa gtg	296
58	Ser Thr Leu Asn Val Ile Val Ser Asn Ser Glu Leu Ala Lys Glu Val	
59	70 75 80	
61	ctc aag gaa aaa gat caa caa ttg gct gat agg cat agg agt aga tca	344
62	Leu Lys Glu Lys Asp Gln Gln Leu Ala Asp Arg His Arg Ser Arg Ser	

RAW SEQUENCE LISTING

DATE: 07/05/2006

PATENT APPLICATION: US/10/681,878B

TIME: 14:22:51

Input Set : A:\44463336.APP

Output Set: N:\CRF4\07052006\J681878B.raw

63	85	90	95	
65	gct gcc aaa ttt agc agg gat ggg cag gac ctt ata tgg gct gat tat	392		
66	Ala Ala Lys Phe Ser Arg Asp Gly Gln Asp Leu Ile Trp Ala Asp Tyr			
67	100 105 110 115			
69	gga cct cac tat gtg aag gtt aca aag gtt tgt acc ctc gag ctt ttt	440		
70	Gly Pro His Tyr Val Lys Val Thr Lys Val Cys Thr Leu Glu Leu Phe			
71	120 125 130			
73	act cca aag cgg ctt gaa gct ctt aga ccc att aga gaa gat gaa gtt	488		
74	Thr Pro Lys Arg Leu Glu Ala Leu Arg Pro Ile Arg Glu Asp Glu Val			
75	135 140 145			
77	aca gcc atg gtt gag tcc att ttt aat gac act gcg aat cct gaa aat	536		
78	Thr Ala Met Val Glu Ser Ile Phe Asn Asp Thr Ala Asn Pro Glu Asn			
79	150 155 160			
81	tat ggg aag agt atg ctg gtg aag aag tat ttg gga gca gta gca ttc	584		
82	Tyr Gly Lys Ser Met Leu Val Lys Lys Tyr Leu Gly Ala Val Ala Phe			
83	165 170 175			
85	aac aac att aca aga ctc gca ttt gga aag cga ttc gtg aat tca gag	632		
86	Asn Asn Ile Thr Arg Leu Ala Phe Gly Lys Arg Phe Val Asn Ser Glu			
87	180 185 190 195			
89	ggt gta atg gac gag caa gga ctt gaa ttt aag gaa att gtg gcc aat	680		
90	Gly Val Met Asp Glu Gln Gly Leu Glu Phe Lys Glu Ile Val Ala Asn			
91	200 205 210			
93	gga ctc aag ctt ggt gcc tca ctt gca atg gct gag cac att cct tgg	728		
94	Gly Leu Lys Leu Gly Ala Ser Leu Ala Met Ala Glu His Ile Pro Trp			
95	215 220 225			
97	ctc cgt tgg atg ttc cca ctt gag gaa ggg gcc ttt gcc aag cat ggg	776		
98	Leu Arg Trp Met Phe Pro Leu Glu Glu Gly Ala Phe Ala Lys His Gly			
99	230 235 240			
101	gca cgt agg gac cga ctt acc aga gct atc atg gaa gag cac aca ata	824		
102	Ala Arg Arg Asp Arg Leu Thr Arg Ala Ile Met Glu Glu His Thr Ile			
103	245 250 255			
105	gcc cgt aaa aag agt ggt gga gcc caa caa cat ttc gtg gat gca ttg	872		
106	Ala Arg Lys Lys Ser Gly Gly Ala Gln Gln His Phe Val Asp Ala Leu			
107	260 265 270 275			
109	ctc acc cta caa gag aaa tat gac ctt agc gag gac act att att ggg	920		
110	Leu Thr Leu Gln Glu Lys Tyr Asp Leu Ser Glu Asp Thr Ile Ile Gly			
111	280 285 290			
113	ctc ctt tgg gat atg atc act gca ggc atg gac aca acc gca atc tct	968		
114	Leu Leu Trp Asp Met Ile Thr Ala Gly Met Asp Thr Thr Ala Ile Ser			
115	295 300 305			
117	gtc gaa tgg gcc atg gcc gag tta att aag aac cca agg gtg caa caa	1016		
118	Val Glu Trp Ala Met Ala Glu Leu Ile Lys Asn Pro Arg Val Gln Gln			
119	310 315 320			
121	aaa gct caa gag gag cta gac aat gta ctt ggg tcc gaa cgt gtc ctg	1064		
122	Lys Ala Gln Glu Glu Leu Asp Asn Val Leu Gly Ser Glu Arg Val Leu			
123	325 330 335			
125	acc gaa ttg gac ttc tca agc ctc cct tat cta caa tgt gta gcc aag	1112		
126	Thr Glu Leu Asp Phe Ser Ser Leu Pro Tyr Leu Gln Cys Val Ala Lys			
127	340 345 350 355			

RAW SEQUENCE LISTING

DATE: 07/05/2006

PATENT APPLICATION: US/10/681,878B

TIME: 14:22:51

Input Set : A:\44463336.APP

Output Set: N:\CRF4\07052006\J681878B.raw

```

129 gag gca cta agg ctg cac cct cca aca cca cta atg ctc cct cat cgc 1160
130 Glu Ala Leu Arg Leu His Pro Pro Thr Pro Leu Met Leu Pro His Arg
131          360          365          370
133 gcc aat gcc aac gtc aaa att ggt ggc tac gac atc cct aag gga tca 1208
134 Ala Asn Ala Asn Val Lys Ile Gly Gly Tyr Asp Ile Pro Lys Gly Ser
135          375          380          385
137 aat gtt cat gta aat gtc tgg gcc gtg gct cgt gat cca gca gtg tgg 1256
138 Asn Val His Val Asn Val Trp Ala Val Ala Arg Asp Pro Ala Val Trp
139          390          395          400
141 cgt gac cca cta gag ttt cga ccg gaa cgg ttc tct gaa gac gat gtc 1304
142 Arg Asp Pro Leu Glu Phe Arg Pro Glu Arg Phe Ser Glu Asp Asp Val
143          405          410          415
145 gac atg aaa ggt cac gat tat agg cta ctg ccg ttt ggt gca ggg agg 1352
146 Asp Met Lys Gly His Asp Tyr Arg Leu Leu Pro Phe Gly Ala Gly Arg
147 420          425          430          435
149 cgt gtt tgc ccc ggt gca caa ctt ggc atc aat ttg gtc aca tcc atg 1400
150 Arg Val Cys Pro Gly Ala Gln Leu Gly Ile Asn Leu Val Thr Ser Met
151          440          445          450
153 atg ggt cac cta ttg cac cat ttc tat tgg agc cct cct aaa ggt gta 1448
154 Met Gly His Leu Leu His His Phe Tyr Trp Ser Pro Pro Lys Gly Val
155          455          460          465
157 aaa cca gag gag att gac atg tca gag aat cca gga ttg gtc acc tac 1496
158 Lys Pro Glu Glu Ile Asp Met Ser Glu Asn Pro Gly Leu Val Thr Tyr
159          470          475          480
161 atg cga acc ccg gtg caa gct gtt ccc act cca agg ctg cct gct cac 1544
162 Met Arg Thr Pro Val Gln Ala Val Pro Thr Pro Arg Leu Pro Ala His
163          485          490          495
165 ttg tac aaa cgt gta gct gtg gat atg taattcttag tttgttatta 1591
166 Leu Tyr Lys Arg Val Ala Val Asp Met
167 500          505
169 ttcattgctct taagggttttg gactttgaac ttatgatgag atttgtaaaaa ttccaagtga 1651
171 tcaaatagaag aaaagaccaa ataaaaaggc ttgacgattt aaaaaaaaaa aaaaaaa 1708
174 <210> SEQ ID NO: 2
175 <211> LENGTH: 508
176 <212> TYPE: PRT
177 <213> ORGANISM: Liquidambar styraciflua
179 <400> SEQUENCE: 2
180 Met Ala Phe Leu Leu Ile Pro Ile Ser Ile Ile Phe Ile Val Leu Ala
181 1          5          10          15
183 Tyr Gln Leu Tyr Gln Arg Leu Arg Phe Lys Leu Pro Pro Gly Pro Arg
184          20          25          30
186 Pro Trp Pro Ile Val Gly Asn Leu Tyr Asp Ile Lys Pro Val Arg Phe
187          35          40          45
189 Arg Cys Phe Ala Glu Trp Ser Gln Ala Tyr Gly Pro Ile Ile Ser Val
190          50          55          60
192 Trp Phe Gly Ser Thr Leu Asn Val Ile Val Ser Asn Ser Glu Leu Ala
193 65          70          75          80
195 Lys Glu Val Leu Lys Glu Lys Asp Gln Gln Leu Ala Asp Arg His Arg
196          85          90          95

```

RAW SEQUENCE LISTING

DATE: 07/05/2006

PATENT APPLICATION: US/10/681,878B

TIME: 14:22:51

Input Set : A:\44463336.APP

Output Set: N:\CRF4\07052006\J681878B.raw

```

198 Ser Arg Ser Ala Ala Lys Phe Ser Arg Asp Gly Gln Asp Leu Ile Trp
199          100          105          110
201 Ala Asp Tyr Gly Pro His Tyr Val Lys Val Thr Lys Val Cys Thr Leu
202          115          120          125
204 Glu Leu Phe Thr Pro Lys Arg Leu Glu Ala Leu Arg Pro Ile Arg Glu
205          130          135          140
207 Asp Glu Val Thr Ala Met Val Glu Ser Ile Phe Asn Asp Thr Ala Asn
208 145          150          155          160
210 Pro Glu Asn Tyr Gly Lys Ser Met Leu Val Lys Lys Tyr Leu Gly Ala
211          165          170          175
213 Val Ala Phe Asn Asn Ile Thr Arg Leu Ala Phe Gly Lys Arg Phe Val
214          180          185          190
216 Asn Ser Glu Gly Val Met Asp Glu Gln Gly Leu Glu Phe Lys Glu Ile
217          195          200          205
219 Val Ala Asn Gly Leu Lys Leu Gly Ala Ser Leu Ala Met Ala Glu His
220          210          215          220
222 Ile Pro Trp Leu Arg Trp Met Phe Pro Leu Glu Gly Ala Phe Ala
223 225          230          235          240
225 Lys His Gly Ala Arg Arg Asp Arg Leu Thr Arg Ala Ile Met Glu Glu
226          245          250          255
228 His Thr Ile Ala Arg Lys Lys Ser Gly Gly Ala Gln Gln His Phe Val
229          260          265          270
231 Asp Ala Leu Leu Thr Leu Gln Glu Lys Tyr Asp Leu Ser Glu Asp Thr
232          275          280          285
234 Ile Ile Gly Leu Leu Trp Asp Met Ile Thr Ala Gly Met Asp Thr Thr
235          290          295          300
237 Ala Ile Ser Val Glu Trp Ala Met Ala Glu Leu Ile Lys Asn Pro Arg
238 305          310          315          320
240 Val Gln Gln Lys Ala Gln Glu Glu Leu Asp Asn Val Leu Gly Ser Glu
241          325          330          335
243 Arg Val Leu Thr Glu Leu Asp Phe Ser Ser Leu Pro Tyr Leu Gln Cys
244          340          345          350
246 Val Ala Lys Glu Ala Leu Arg Leu His Pro Pro Thr Pro Leu Met Leu
247          355          360          365
249 Pro His Arg Ala Asn Ala Asn Val Lys Ile Gly Gly Tyr Asp Ile Pro
250          370          375          380
252 Lys Gly Ser Asn Val His Val Asn Val Trp Ala Val Ala Arg Asp Pro
253 385          390          395          400
255 Ala Val Trp Arg Asp Pro Leu Glu Phe Arg Pro Glu Arg Phe Ser Glu
256          405          410          415
258 Asp Asp Val Asp Met Lys Gly His Asp Tyr Arg Leu Leu Pro Phe Gly
259          420          425          430
261 Ala Gly Arg Arg Val Cys Pro Gly Ala Gln Leu Gly Ile Asn Leu Val
262          435          440          445
264 Thr Ser Met Met Gly His Leu Leu His His Phe Tyr Trp Ser Pro Pro
265          450          455          460
267 Lys Gly Val Lys Pro Glu Glu Ile Asp Met Ser Glu Asn Pro Gly Leu
268 465          470          475          480
270 Val Thr Tyr Met Arg Thr Pro Val Gln Ala Val Pro Thr Pro Arg Leu

```

RAW SEQUENCE LISTING

DATE: 07/05/2006

PATENT APPLICATION: US/10/681,878B

TIME: 14:22:51

Input Set : A:\44463336.APP

Output Set: N:\CRF4\07052006\J681878B.raw

```

271          485          490          495
273 Pro Ala His Leu Tyr Lys Arg Val Ala Val Asp Met
274          500          505
278 <210> SEQ ID NO: 3
279 <211> LENGTH: 1883
280 <212> TYPE: DNA
281 <213> ORGANISM: Liquidambar styraciflua
283 <220> FEATURE:
284 <221> NAME/KEY: CDS
285 <222> LOCATION: (74)..(1606)
287 <400> SEQUENCE: 3
288 tgcaaacctg cacaacacaa gagagagaag aagaaaaagg aagagaggag agagagagag 60
290 agagagagaa gcc atg gat tct tct ctt cat gaa gcc ttg caa cca cta 109
291          Met Asp Ser Ser Leu His Glu Ala Leu Gln Pro Leu
292          1          5          10
294 ccc atg acg ctg ttc ttc att ata cct ttg cta ctc tta ttg ggc cta 157
295 Pro Met Thr Leu Phe Phe Ile Ile Pro Leu Leu Leu Leu Gly Leu
296          15          20          25
298 gta tct cgg ctt cgc cag aga cta cca tac cca cca ggc cca aaa ggc 205
299 Val Ser Arg Leu Arg Gln Arg Leu Pro Tyr Pro Pro Gly Pro Lys Gly
300          30          35          40
302 tta ccg gtg atc gga aac atg ctc atg atg gat caa ctc act cac cga 253
303 Leu Pro Val Ile Gly Asn Met Leu Met Met Asp Gln Leu Thr His Arg
304          45          50          55          60
306 gga ctc gcc aaa ctc gcc aaa caa tac ggc ggt cta ttc cac ctc aag 301
307 Gly Leu Ala Lys Leu Ala Lys Gln Tyr Gly Gly Leu Phe His Leu Lys
308          65          70          75
310 atg gga ttc tta cac atg gtg gcc gtt tcc aca ccc gac atg gct cgc 349
311 Met Gly Phe Leu His Met Val Ala Val Ser Thr Pro Asp Met Ala Arg
312          80          85          90
314 caa gtc ctt caa gtc caa gac aac atc ttc tcg aac cgg cca gcc acc 397
315 Gln Val Leu Gln Val Gln Asp Asn Ile Phe Ser Asn Arg Pro Ala Thr
316          95          100          105
318 ata gcc atc agc tac ctc acc tat gac cga gcc gac atg gcc ttc gct 445
319 Ile Ala Ile Ser Tyr Leu Thr Tyr Asp Arg Ala Asp Met Ala Phe Ala
320          110          115          120
322 cac tac ggc ccg ttt tgg cgt cag atg cgt aaa ctc tgc gtc atg aaa 493
323 His Tyr Gly Pro Phe Trp Arg Gln Met Arg Lys Leu Cys Val Met Lys
324          125          130          135          140
326 tta ttt agc cgg aaa cga gcc gag tcg tgg gag tcg gtc cga gac gag 541
327 Leu Phe Ser Arg Lys Arg Ala Glu Ser Trp Glu Ser Val Arg Asp Glu
328          145          150          155
330 gtc gac tcg gca gta cga gtg gtc gcg tcc aat att ggg tcg acg gtg 589
331 Val Asp Ser Ala Val Arg Val Val Ala Ser Asn Ile Gly Ser Thr Val
332          160          165          170
334 aat atc ggc gag ctg gtt ttt gct ctg acg aag aat att act tac agg 637
335 Asn Ile Gly Glu Leu Val Phe Ala Leu Thr Lys Asn Ile Thr Tyr Arg
336          175          180          185
338 gcg gct ttt ggg acg atc tcg cat gag gac cag gac gag ttc gtg gcc 685

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/05/2006
PATENT APPLICATION: US/10/681,878B TIME: 14:22:52

Input Set : A:\44463336.APP
Output Set: N:\CRF4\07052006\J681878B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:21; Xaa Pos. 3,5,6,8

Seq#:22; N Pos. 23

VERIFICATION SUMMARY

DATE: 07/05/2006

PATENT APPLICATION: US/10/681,878B

TIME: 14:22:52

Input Set : A:\44463336.APP

Output Set: N:\CRF4\07052006\J681878B.raw

L:1259 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0

L:1278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0